Proceedings of the Sixteenth International AAAI Conference on Web and Social Media (ICWSM 2022)

Time after Time: Longitudinal Trends in Nostalgic Listening

Clara Hanson,[1*] Jesse Anderton,[2] Samuel F. Way,[2] Ian Anderson,[2] Scott Wolf,[2] Alice Wang[2]

[1] University of California, Los Angeles
[2] Spotify
clarahanson@ucla.edu

Abstract

Nostalgia was once considered a medical disease but is now understood as a beneficial, identity-affirming emotion. Yet what induces feelings of nostalgia is not fully understood. Are nostalgic feelings prompted by changes in a person's life, and do certain events prompt nostalgia across whole populations? In this paper, we analyze when people listen to nostalgic music. We use data from a large, cross-national survey to train a classifier to detect which tracks are nostalgic for individual listeners. We then analyze a comprehensive dataset of listening histories from a 5.5 year period. Despite it being a complex concept, we were able to predict nostalgic listening with relatively high precision. We compare our results across listeners from four countries to understand how consistent behavior toward nostalgic music is. We find people listen to nostalgic music more often as they age. We also find people tend to listen to nostalgic music consistently in their day-to-day lives. We do not find evidence that listening to personally nostalgic music increases in response to particular traditions, seasons, or events. However, we do find traditions and events can affect how much "back catalog" music people listen to. These trends are consistent across the national contexts we studied. Our results advance prior findings about nostalgia and the life course, and demonstrate a novel methodological consideration for studies of nostalgic listening.

Introduction

Nostalgia has many meanings. At times, it refers to consumer products or marketing that reproduce trends from the past (Merchant and Rose 2013). At other times, it refers to the romanticization of history as a political tool (González 2012). Work in psychology understands nostalgia as an emotion of sentimental longing felt toward a person's own memories, life events, and relationships (Davis 1979; Sedikides and Wildschut 2018). This definition focuses on lived, personal memories, in contrast to concepts of nostalgia that focus on stories told by groups of people, which may include periods beyond their lived memory.

One may expect a person's experience of time to trigger nostalgic emotion. These include life course events like birthdays, graduations, weddings, and funerals. It may also include more gradual processes, like aging. Taste in music also changes with age (Bonneville-Roussy et al. 2013), and prior research suggests that people become more nostalgic with age (Holbrook 1993) and recall memories more positively with age (Schlagman et al. 2009). Yet other studies suggest people often feel music-evoked nostalgia in daily life (Jakubowski and Ghosh 2021).

Collective experiences of time — like annual traditions or public events — could also trigger nostalgic emotion. Prior work finds the emotional tone of music people listen to tends to vary over annual cycles (Park et al. 2019). Recently, researchers have also identified an increase in listening to nostalgic music following the onset of the Covid-19 pandemic (Yeung 2020). This is consistent with findings that people use nostalgia to cope with hardship (Sedikides, Wildschut, and Baden 2004). Both suggest social and environmental factors affect the emotions people feel.

Understanding when people feel nostalgic contextualizes feelings and behaviors related to identity, mental health, and social connectedness. Research across the social sciences documents the positive psychological and prosocial effects of nostalgic emotion (Routledge et al. 2011; Hepper et al. 2021; van Tilburg, Igou, and Sedikides 2013; Zhou et al. 2011; Juhl et al. 2021). Notably, listening to nostalgic music strengthens a person's sense of continuity, meaning, and identity in their life (Sedikides and Wildschut 2018). This contrasts older notions of nostalgia, which are more negative (Batcho 2013; De Diego, Ots et al. 2014).

Despite these insights, understanding how experiences of nostalgic emotion unfold in time is a difficult task. First, it is difficult because there are many different reasons a person may listen to old music: to practice traditions, explore social or musical history, or simply to hear music they like (DeNora 2000). Music from the "back catalog" may have any number of symbolic meanings that are difficult to distinguish without more information about the listener's relationship to it. Second, observing a person's behavior over a long period is difficult, and becomes even more difficult at a large scale. However, diary studies like Jakubowski and Ghosh (2021) demonstrate the importance of observing behavior in everyday life to our understanding of music, memory, and emotion.

Present Study

The primary aim of this study is to understand how listening to nostalgic music is related to personal and collective senses of time. By "personal senses of time," we mean temporal patterns that must be controlled for at the level of the individual (e.g., age). By "collective senses of time," we mean temporal patterns that affect large groups at the same time (e.g., holidays, seasons). Nostalgic emotion is a sentimental longing for one's past, and so we expect nostalgia to relate to personal senses of time. However, longing for the past may also be a reaction to events in the present. Therefore, we also expect changes in nostalgia across large groups of people.

The data we use in this study is uniquely suited to analyze the relationship between nostalgic listening and time. Our data is drawn from Spotify, a large music streaming service. First, we conducted a large, cross-national survey in which participants labeled songs from their own listening histories. We trained a classifier to predict whether particular songs are nostalgic for particular listeners. We then collected a random sample of 400k listeners from Brazil, Spain, the United Kingdom, and the United States. We applied the classifier to the music they listened to between 2016 to mid-2021, and analyze the predictions of that classifier for this sample.

Historically, studies about music, memory, and emotion are unable to sample at this scale and level of specificity. This work overcomes three common limitations among existing studies: (1) representation across age cohorts, (2) representation across national contexts, and (3) the ability to analyze patterns at the level of individuals, at scale.

For each result, we compare nostalgic listening to catalog listening. This approach allows us to consider whether people stream nostalgic music any differently than the way they stream old music more generally. Studies of music, memory, and emotion are rarely able to study participants in such a personalized way. Comparing these distributions also allows us to consider whether the patterns observed with a less personalized approach to classifying nostalgic music capture the same general distribution as more personalized methods. This analysis provides additional context about the findings and limitations of studies based on aggregate listening statistics or samples of tracks selected by researchers.

For each analysis, we show how our results vary by national context. Hepper et al. (2014)'s analysis shows nostalgia exists in a wide range of cultural contexts. However, additional work suggests the particular meaning of nostalgia varies across cultures (Farese and Asano-Cavanagh 2019). Some evidence suggests tendency toward nostalgia varies by a country's collectivism or individualism (Abakoumkin, Wildschut, and Sedikides 2020; Granot et al. 2021). We consider how robust trends are across national contexts to assess whether cultural differences result in different nostalgic listening behavior.

To summarize, there are three central aims of our study. First, we use a novel dataset to examine how nostalgic listening is affected by temporal factors. Second, we consider whether nostalgic listening is similar or different across cultures. Third, we evaluate whether patterns in nostalgic listening are captured by catalog listening more generally.

Data and Methods

Our study analyzes how listeners interact with nostalgic music. In the section that follows, we outline three components that make this analysis possible. First, we ran a large, cross-national survey. In the survey, listeners indicated music from their listening histories they found nostalgic. Second, we developed a random forest classifier to reproduce the answers of these survey responses to accurately label instances of nostalgic listening. Finally, we sampled listeners and collected their comprehensive listening histories. We applied our classifier to this data, and analyzed patterns of nostalgic listening over time and across national contexts.

A survey to identify personally nostalgic tracks. We expected nostalgic listening to make up a small percentage of the music people listen to. Randomly sampling tracks for respondents to label would yield too few positive responses to train an accurate classifier, and so we needed to a method to over-sample nostalgic tracks. To do so, we surveyed a convenience sample of about 500 employees of Spotify based in the United States and Sweden. This ensured high-quality responses and allowed us to iteratively test hypotheses. In the final employee survey, respondents labeled tracks that varied in affinity, release year, and age-cohort over-index score. With this data, we trained a classifier to score the probability that a given listener would find a track nostalgic. We restricted the number of features used by the classifier to avoid overfitting. We applied the classifier to respondents' listening histories to sample tracks for the listener survey. Among the tracks we asked respondents to label, 50% had high probabilities of being nostalgic. The other 50% had low probabilities of being nostalgic, but were likely to be recognized by respondents (e.g., because the respondent had affinity for the track). We did this to better identify the boundary between nostalgic and non-nostalgic tracks. This classifier was more accurate at predicting nostalgic tracks for respondents in the United States than in other countries. This underscores the bias in this sampling method. Employees in this sample differ from the global population in ways that could affect our results (e.g., tend to be Western, young, have high levels of education, etc.). Thus, the likely-nostalgic tracks we over-sampled to create our survey may better reflect concepts of nostalgia held by Spotify employees than those of the global population. However, that we asked respondents to label these tracks helps to mitigate some of this bias.

We ran the listener survey in 2017. All tracks included in the survey were released in 2013 or earlier. Respondents labeled only tracks released after they were born. This aligns with the idea that nostalgic emotion is about personal, lived experiences. However, it may exclude nostalgic music released before the respondent's birth, or nostalgic music the respondent has not listened to on Spotify.

To survey Spotify listeners about what tracks they find nostalgic, we invited a random sample of listeners to participate via email. Responses were collected in August of 2017. In the survey, respondents were shown up to 12 tracks, and labeled them as either "nostalgic" or "not nostalgic." Track name, artist, and album art were displayed to respondents in the survey. Respondents in this analysis were between age 18 and 67 at the time of the survey (Figure 1). The survey was written in English and then translated into the official languages of each country surveyed. Respondents were located in 11 countries. We surveyed a total of 17,547 re-

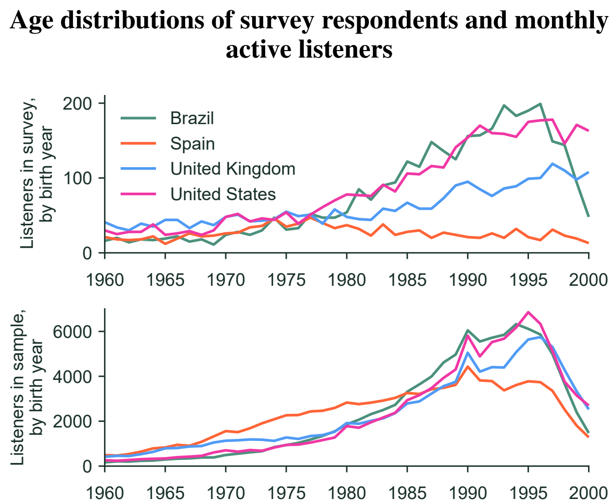

Figure 1: Age distributions of survey respondents (top) and the sample of monthly active listeners (bottom) in the study.

spondents who labeled a total of 196,011 tracks. Respondents labeled 62.7% of the tracks nostalgic, suggesting our method underestimated the decision boundary between nostalgic and non-nostalgic tracks. For more information about the sample, see Appendix Table A1.

We did not define nostalgia in the survey. We believe this approach is appropriate for the aims of this analysis. The purpose of surveying such a diverse sample of listeners was to understand what different people think nostalgia is, and what kinds of music they feel nostalgic toward. There is no definitive cross-cultural definition of nostalgia; it has different connotations across cultures. Offering a definition could bias respondents and reproduce our assumptions of what nostalgia is. Thus, the labels we collected have validity because they reflect listeners' own concepts of nostalgia.

A classifier to predict nostalgic tracks. Our labeled data over-represents nostalgic and nostalgic-seeming music. To better account for the breadth of music people listen to, we randomly sampled tracks from respondents' listening histories and added them to the training data as negatively labeled data. The number of assumed-negative tracks we added was equal to the number of labeled data points for each respondent. Although this may introduce false negatives to the dataset, we expect the proportion to be negligible for most listeners (for further discussion, see Appendix Figure A1). The proportion of false negatives is likely highest among the youngest listeners, but our final classifier's recall was highest among young listeners born in the early 1990s and did not drop substantially for even younger respondents (see Figure 4). We found training the classifier with assumed-negative data resulted in a classifier with a higher AUC. Introducing false negatives to the dataset may decrease recall, but we judged improving the classifier's precision a higher priority.

We trained and evaluated a random forest classifier using this data. First, we created train, test, and validation datasets by randomly partitioning respondents. This helps reduce overfitting because the classifier was not trained using data from respondents in the test or validation datasets. 60% were assigned to the training dataset, and 20% each were assigned to the testing and validation datasets. We stratified the data by age to train the classifier. We tried two approaches to account for cross-national differences in our data: training classifiers based on each country's data, and training a single classifier with listener country included as a feature. We found the second approach achieved higher precision than the first approach on the testing data. Thus, we developed a single classifier and used data from respondents in all countries.

The classifier we developed uses 45 features to predict whether a track is nostalgic for each listener. The features we use include: listeners' self-reported demographic information (e.g., birth year, gender, country), temporal information (e.g., track release year, listener age when the track was released), popularity information (e.g., how many times the song or artist has been streamed on Spotify), audio and genre characteristics (i.e., top-level Gracenote genre, Echo Nest acoustic features), information about how a listener interacts with music (e.g., a measure of the listener's affinity for a track, artist, or genre based on their streaming history), information about how others the listener's age or from the listener's country relate to music (e.g., a z-score for how much others the listener's age play the track compared to the population mean), and information about the seasonality of streams (e.g., entropy of number of streams over weeks of the year).

Features related to time and interaction were most useful for predicting nostalgic music (Figure 2). The year a track was released and the listener's age at that time were informative time-related features. Appendix Figure A2 further explores how our classifier represents temporal features. Features related to interaction are less straightforward. They span how listeners themselves interact with tracks (e.g., listener affinity), how others in the listener's age group interact with tracks (e.g., cohort streamshare and overindexing scores), and general popularity of the track (e.g., stream counts). Musical attributes are less informative (e.g., genre, Echo Nest acoustic features). This suggests nostalgia is difficult to forecast, and becomes easier to recognize as people revisit the artifacts that become nostalgic to them.

Our final model is a random forest classifier that predicts whether a track is nostalgic for a particular listener. We chose this approach after evaluating several different approaches because it has both high predictive performance and is relatively interpretable. Our final model uses an ensemble of 100 estimators with no maximum depth, and uses Gini impurity to measure the quality of branching splits.

We tuned the model on the testing data, and chose a threshold that favored precision first and recall second. The model achieved a precision of 0.71 and a recall of 0.48 on the validation data. Performance was higher on average in the US and Brazil, and lower on average in Spain (Figure 3). We believe lower performance for respondents in Spain is due to different connotations of English "nostalgia" and Spanish "nostálgica" (see Appendix Figure A3). Precision was simi-

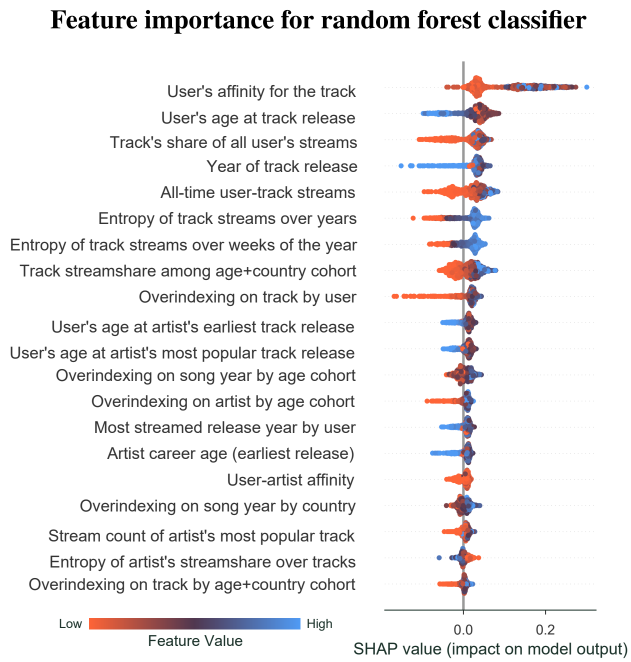

Figure 2: Overview of which features most affected our classifier's predictions. The twenty most informative features are ordered from most informative (top) to least (bottom). Dots show SHAP value of each feature for a random user-track pair. The color corresponds to the value of the feature, where blue indicates a high value and orange indicates a low value.

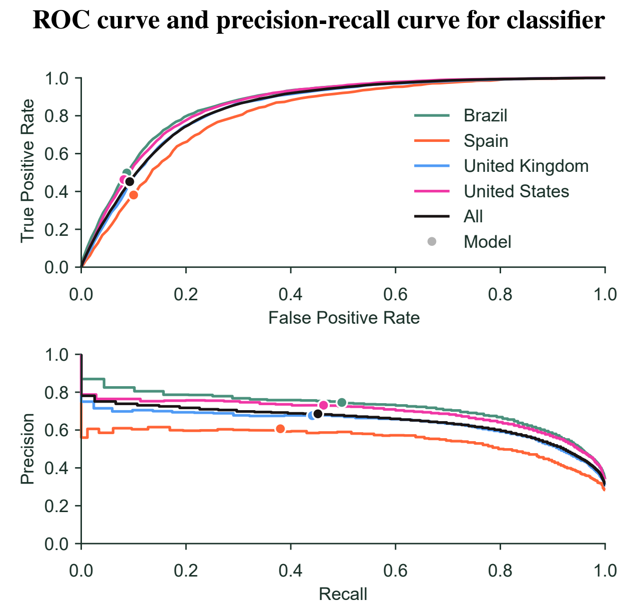

Figure 3: ROC curve (top) and precision-recall curve (bottom) of our classifier. These were computed using the validation data, which contains both labeled data and an equal number of assumed-negative tracks. We evaluated model performance with the validation data only after selecting the classifier used in this paper. Points on each line indicate model performance at the model's classification threshold.

lar across track release years and listener birth years, meaning the chance a predicted track truly is nostalgic does not vary by the age of a listener or era of music. However, recall was higher for younger respondents (Figure 4). We consider how this could affect our results where relevant.

However imperfect, the degree to which our classifier predicts what respondents found nostalgic is noteworthy. Respondents labeled tracks using their own ideas of nostalgia, and while nostalgia is a complex emotion with different connotations, there is significant predictability in what people find nostalgic. That said, our classifier may only allow us to study instances of nostalgic listening that are most easily predicted.

Applying the classifier to listening histories Finally, we created a large, longitudinal dataset to analyze nostalgic listening over time. This dataset is composed of a random sample of approximately 100,000 listeners from four of the eleven countries represented in the training data: Brazil, Spain, the United Kingdom (UK), and the United States (US). The longitudinal dataset contains all instances where a listener streamed a track for 30+ seconds between January 2016 and July 2021. We limited the sample to listeners who streamed at least one song per month for the entire 5.5 year sample frame (Figure 1). This reduces the chance of including listeners who do not use Spotify as their primary means of listening to music. This does not create selection bias among Spotify's listeners in general (Sanna Passino et al. 2021). As a check, we compared the track ages and genres streamed in this dataset to one in which users did not listen to at least one track every month. We found no notable differences between the two. We applied our classifier to this dataset.

Though the survey was conducted in eleven countries, we focus on listeners in only four. We do this for two reasons. First, for practical reasons. We focus on countries where we received a high number of responses to the listener survey, and where Spotify operated for long enough to collect a large, longitudinal sample. Second, for analytical reasons. We chose countries that maximize the diversity of languages and cultural contexts. In particular, we wanted to assure our analysis included countries with tendencies toward individualism (e.g., the US) and tendencies toward collectivism (e.g., Brazil) (Hofstede 1984; Triandis 1995).

Throughout our analysis, we compare nostalgic listening to "catalog listening." We define catalog listening as streaming any track released in 2013 or earlier. This category is intentionally broad, because it allows us to compare nostalgic streaming to listening to older music more generally. We use 2013 as the threshold because 2013 was the most recent release year of tracks included in the listener survey.

Two methods we use in our analysis require elaboration. First, we use Prophet (Taylor and Letham 2018), a regression framework similar to generalized additive models (GAMs), to model time trends in nostalgic listening. Prophet was designed to model seasonalities common in business time-series data, including regressors for time of year, day

Precision and recall by respondent age and track release year

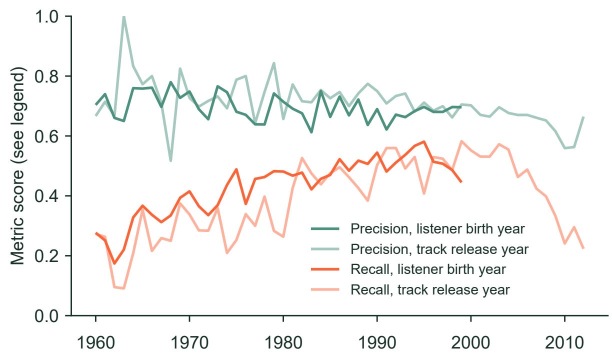

Figure 4: The figure above shows out classifier's precision and recall by respondent birth year and track release year. Precision varies across years but does not show a linear trend. Recall increases over years for both graphs, and decreases for the most recent tracks.

Tendency to stream nostalgic music by age group

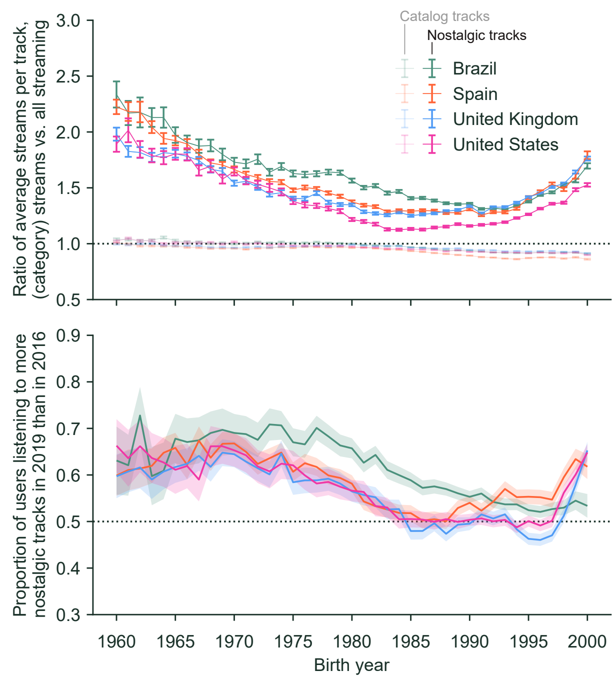

Figure 5: Top: Number of streams to nostalgic tracks (bold), and number of streams to catalog tracks (faded) by birth year, normalized for total amount of streaming. Error bars show standard error of the mean. Bottom: Proportion of listeners in each age group who streamed more nostalgic tracks in 2019 than they did in 2016. Shading shows 95% confidence interval. Track affinity was an important feature of the classifier, and so higher rates of nostalgic streaming compared to catalog streaming is not surprising.

of week, and country-specific holidays. Each of these seasonalities significantly shape listening on Spotify and are interesting in their own right. Separating them from overall trends affords interpretation of how nostalgic listening has evolved over time and in response to key events.

Second, we measure how much listeners stream the same nostalgic music. To measure similarity, we computed tf-idf weighted matrices based on the number of times each listener streamed each track. We excluded tracks streamed by only one person in each country. Each matrix was normalized using the Euclidean norm so we could measure similarity between listeners with cosine similarity. We compare similarity across age groups by calculating the average of all similarity scores between all pairs of listeners from two birth years. We measure similarity within age groups by calculating the same measure, but subtracting each listener's self-similarity. Similarity by nostalgic listening compares all tracks labeled nostalgic (a total of 45,375 unique tracks and 42,471,434 listener-track pairs). Similarity by catalog listening compares a random sample of catalog tracks (a total of 181,135 unique tracks and 53,032,715 listener-track pairs).

Results

We are interested in understanding how nostalgic listening relates to time. First we consider temporal processes and events that are experienced on a personal scale (e.g., aging, life course events). Then we consider events and processes that affect much larger groups (e.g., annual events, public events). Finally, we explore how people relate to their cohort and society. In addition to evaluating general trends in nostalgic listening, we compare each result to catalog listening generally. We show results by country throughout our analysis to consider how consistent nostalgic listening is across different cultural contexts.

Personal time Nostalgia can affirm a person's sense of identity, and may be used as a resource when coping with events and changes in a person's life.

We find people listen to more nostalgic music with age (Figure 5, top). These estimates are normalized by how much music each listener streams. Beginning with listeners born in the early 1980s, the curve shows listeners stream more nostalgic music with age. Streams to catalog tracks also increase with age, but the size of the increase is very small. These trends are consistent across national contexts.

The youngest listeners in our sample stream nostalgic music at a higher rate than listeners 5- to 10-years older. This is likely because nostalgic music for young listeners is still relatively modern. Modern music and nostalgic music only become distinct as a listener ages.

Our data suggest increases in nostalgic listening across age groups is in part an effect of aging (Figure 5, bottom). Listeners born in about 1980 or earlier listened to more nostalgic music in 2019 than they did in 2016. Younger listeners in the US and UK did not show the same trend. Most listeners in Brazil listened to more nostalgic music with time, but the proportion is higher among older listeners. We compare 2016 to 2019 because 2016 is our earliest year of data, and

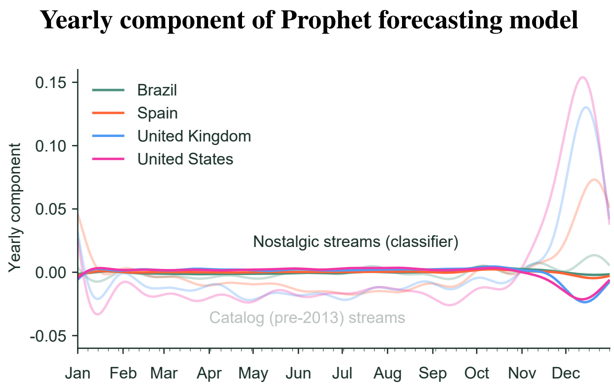

Figure 6: The yearly components of our Prophet forecast models for catalog streaming (pastel), and personally nostalgic streaming (bold). Data stratified by birth year.

because the Covid-19 pandemic complicates a direct comparison to years after 2019.

This analysis also helps us consider whether differences in nostalgic listening across age groups are an artifact of our methods. The classifier we used had lower recall for older listeners, and its predictions are partly based on a listener's affinity for a track. If the threshold for track affinity was higher for older listeners than for younger listeners, this could create an apparent increase in nostalgic listening across age groups. However, we observe that a majority of listeners born in the early 1980s or earlier listened to more nostalgic music over time. This provides additional support that the rate of nostalgic listening increases with age.

Another alternative explanation for these trends is increases in Spotify's catalog over this period. However, Spotify's catalog of popular tracks is quite complete. For each year from 1960–2013, the catalog contains at least 89 of the Billboard Hot 100 tracks. It contains 96.5% of all tracks for the same period. The majority of Billboard Hot 100 tracks available in 2020 were also available in 2016, though more tracks were added from earlier years of music than from later years. While we cannot rule out some effect of the catalog on nostalgic listening, a great deal of catalog music was available to listeners throughout the period we analyzed. We do not expect it would change the trend of these findings.

We are also interested in whether people listen to nostalgic music in response to life course events (e.g. birthdays, weddings). We do not have information about when these events occurred for each respondent. However, we expected that if people listen to nostalgic music in response to life course events, nostalgic music would be clustered in time rather than consistent over the observed period. We calculated entropy to explore how consistently respondents streamed nostalgic music over weeks of the sample frame. We found listeners had high entropy scores on average, indicating listeners stream nostalgic music relatively consistently. Younger listeners had relatively higher entropy scores, and older listeners had relatively lower entropy scores. This difference may be caused by the lower recall of our classifier for older respondents, and so further

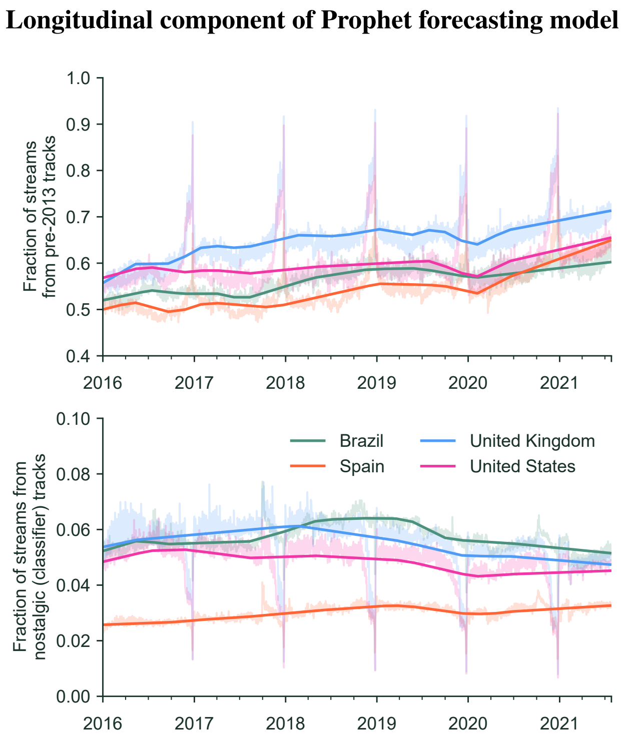

Figure 7: The plots above show the longitudinal components of our Prophet forecast models for catalog streaming (top), and personally nostalgic streaming (bottom). Raw data shown in pastel, the forecast component shown in bold. Data stratified by birth year.

analysis is needed to understand whether nostalgic listening does become more sporadic as people age. See Appendix Figure A4 for distributions.

Annual events and societal change Listening to music can be a social resource that increases cohesion with others. In this section, we explore whether people shift their listening habits en masse in response to annual or public events. The results demonstrate the importance of how "nostalgia" is operationalized in research.

First, we explore whether nostalgic listening follows yearly trends (Figure 6). Nostalgic streaming is relatively consistent throughout the year, though there is a small dip in nostalgic streaming in in the US and UK in December. In contrast, catalog listening spikes beginning in November in the US, UK, and Spain. At its peak, catalog streaming accounts for about 90% of streams annually during this period in the US and UK (Figure 7). Most of the songs that account for this change are about Christmas, but some are about Hanukkah, New Years, and winter generally. Decreases in streams of nostalgic music during this period is likely due to holiday music listening "crowding out" nostalgic listening. Catalog streaming also spikes in Brazil during December, but this spike is mostly isolated to Christmas day (Figure 7). The extent to which people stream holiday music appears to be associated with national context, but not necessarily with the observation of Christmas itself. We find no evidence of

Similarity in catalog and nostalgic listening by birth year

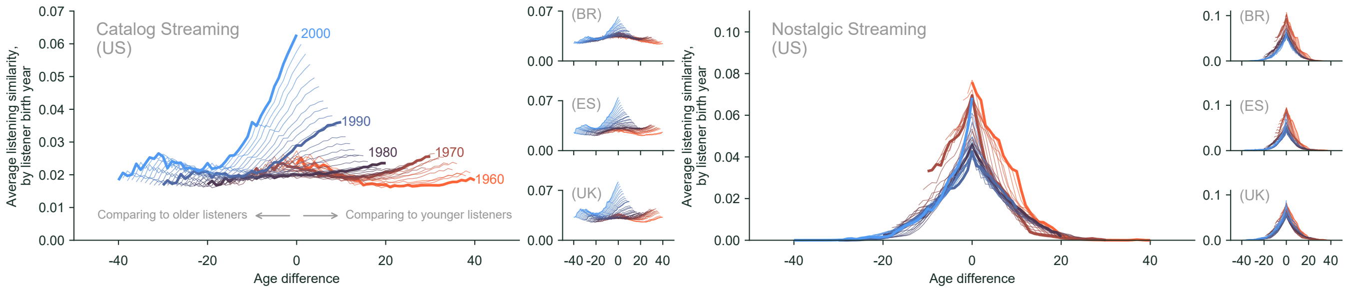

Figure 8: Average similarity among people across age groups by catalog streaming (left) and nostalgic streaming (right). The largest frames show similarities among listeners in the US. The inset images show distributions for Brazil, Spain, and the UK. The y-axis shows average cosine similarity measured using a Euclidean normed tf-idf matrix.

other annual periods where people consume catalog music to the same extent.

We also consider whether events or non-annual cycles affect catalog and nostalgic streaming (Figure 7). During the sample frame, catalog streaming increased somewhat gradually across the four countries. There are no apparent cross-national trends for nostalgic listening. However, beginning in early 2020, there is a meaningful increase in catalog streaming among listeners in the US, UK, and Spain. This suggests people did begin to stream more catalog music following the onset of the Covid-19 pandemic, but that this was not driven by listening to personally nostalgic music. We did not find evidence this increase was driven by a decrease in the amount of music released in 2020. The number of new tracks released in 2020 was higher than those released in 2019. In addition, we found no evidence listeners discovered fewer tracks in 2020 than in 2019.

To summarize, our data suggest streams to catalog music did increase following the Covid-19 pandemic. However, we do not find evidence this increase was driven by personally nostalgic music. This difference suggests studies of listening behavior should account for the symbolic resonance of music to explain changes in listening behavior. Music in the back catalog has many different symbolic meanings that produce different temporal patterns.

The individual and the aggregate Finally, we examine how individual people relate to their society more generally. We do this in two ways. First, we explore similarity in music listening within and across age cohorts. Second, we test whether the tendency to participate in cultural music practices is associated with nostalgic streaming.

Perhaps unsurprisingly, people born around the same time tend to have more nostalgic streaming in common (Figure 8, right). Older listeners in the sample are the most similar to their own age group (orange), and listeners born in the 1980s and early 1990s are the least (darkest blue). The oldest listeners in our sample also tend to have more nostalgic music in common with listeners from a wider range of birth years.

Though this observation may suggest people become more similar to their own cohort with age, we find no evidence that within-cohort similarity changed between 2016–2019. We calculated the cosine similarity for all pairs of listeners within an age group and took the average for each year in the dataset. The averages varied minimally from year to year, and we did not observe consistent longitudinal trends. It is not clear from our data what causes the differences in similarity we observe across birth years. The four year period we consider may be too small to detect longitudinal trends in similarity. Alternatively, these patterns may be a cohort effect. Yet another explanation is that differences in our classifier's recall across age groups creates the appearance that older listeners are more similar than they are. For this reason, it is difficult to assess whether the cross-cohort differences in similarity suggested by this analysis are real or an artifact of our method.

How similar people the same age are in the catalog music they stream follows a very different distribution (Figure 8, left). Within-age group similarity is negatively associated with age. The youngest listeners have the most music in common. The negative association between similarity and age reverses among the oldest listeners, but their overall similarity is still much lower compared to the youngest listeners.

This analysis also suggests an increase in similarity among parent-child groups. Listeners born in the 1990s are more similar to listeners born about 30 years earlier than they are to listeners born 10–20 years earlier. Conversely, listeners born in the 1970s are more similar to listeners born 30 years later than they are to listeners born 10–20 years later. A possible explanation for this trend is parents and children streaming the same music.

The trends in catalog streaming in the US are similar to those observed in other countries. However, listeners in Brazil have higher similarity overall, and similarity scores follow a more convex shape around others close in age. The relationship between birth cohort and catalog streaming varies somewhat by national context.

Finally, we tested whether a person's tendency to listen to personally nostalgic music is associated with their degree of participation in collective streaming events. We operationalize "collective streaming events" by identifying tracks that have relatively high entropy over years (>1.25 nats), but relatively low entropy over weeks in the year (<3 nats). This identifies tracks that are consistently played year after year, but only on certain weeks. We calculated Pearson correlations for the relationship between amount of nostalgic streaming and amount of collective streaming for listeners within each country. All correlations were very weak (ranging from 0.04–0.18). Thus, we did not find evidence that a person's tendency to participate in collective listening and tendency to listen to personally nostalgic music are related.

Discussion

In our analysis, we studied how nostalgic emotion relates to time in several different ways. First, we focused on how nostalgia relates to personal senses of time. We find that people listen to more nostalgic music as they age. This is consistent with past research (Holbrook 1993). We also find people listen to nostalgic music relatively consistently over time, affirming that nostalgia is a common experience when listening to music (Jakubowski and Ghosh 2021). Though this suggests nostalgic listening tends to be habitual, further research can explore whether nostalgic listening becomes a more sporadic practice with age.

We did not account for play context in our analysis (i.e., whether a nostalgic song is played from search, a user's library, a playlist, etc.). Future research can consider whether temporal patterns of nostalgic listening depend on whether a listener actively searches for nostalgic tracks. Incorporating information about listener context can also better account for the effects of habitual listening — compared to socializing and events — on when listeners turn to nostalgic music. Finally, additional research can explain what causes people to listen to more nostalgic music as they age. For example, whether it is caused by changes in the way people use music as they age (Hird and North 2021) or changes in a person's relationship to their identity (Roy and Dowd 2010).

Next, we focused on collective experiences of time. We did not find evidence of cycles or events that caused increases in nostalgic listening at the population level. However, we found annual cycles and events associated with catalog streaming.

First, catalog streaming increases as people stream holiday music at the end of the year. In the US and UK, people treat the end of the year as the Christmas season by listening to holiday music well before Christmas Day. This is true in Spain to a lesser extent. In contrast, holiday music is streamed primarily on Christmas Day in Brazil. In our sample, no other tradition or season compels such a dramatic shift away from contemporary music. This finding is a reminder that some old tracks have symbolic meaning that is very different from personal nostalgic emotion. It also illustrates the role of collective identity and cultural context in shaping traditions (e.g. Weinberger 2015).

We found catalog streaming increased following the onset of the Covid-19 pandemic. Prior research found music was an important resource people used to cope with the sudden difficulty of the global spread of Covid-19 and subsequent stay-home policies (Vidas et al. 2021; Fink et al. 2021; Gibbs and Egermann 2021; Cabedo-Mas, Arriaga-Sanz, and Moliner-Miravet 2020; Granot et al. 2021). In particular, Yeung (2020) found listening to catalog music increased during lockdown periods, and proposed this was driven by an increase in listening to nostalgic music. Our study builds on this by comparing catalog streaming to songs we estimate are personally nostalgic for individual listeners. Our research affirms that listeners' streaming habits changed in response to the pandemic. However, it suggests that a different meaning — not nostalgia — is responsible for this change.

Nostalgia can help people cope with difficult life events (Sedikides, Wildschut, and Baden 2004). However, it can also increase perception of death threat (Yang et al. 2021). More research is necessary to explain what prompts people to listen to nostalgic music, and what drove people to catalog music in response to the Covid-19 pandemic. For example, whether the unique collective nature of the hardship of the Covid-19 pandemic encouraged people to revisit popular, rather than personally nostalgic, music.

Throughout our analysis, we compared personal nostalgic listening to catalog listening. The differences we find underscore that the way researchers operationalize "nostalgia" can result in very different findings. Nostalgia is related to a person's self-image and narrative (Sedikides and Wildschut 2018). Catalog music is a very broad category, and there are many distinct meanings people associate with songs. A person's affinity for a track was an important predictor of nostalgia. Even though people close in age tend to be the most similar in their nostalgic taste, their overall similarity is relatively low. Together, these findings underscore how difficult it is for researchers to know if tracks will hold particular resonance for listeners when sampling a catalog for surveys and experiments. Still, understanding the meaning listeners associate with songs is important for research. Despite the association between nostalgia and collectivity (Abakoumkin, Wildschut, and Sedikides 2020), we found no association between participation in periods of collective catalog streaming and nostalgic streaming at the individual level. The social mechanisms that drive participation in social rituals may be different from those related to more personal practices.

Though this research focuses on nostalgic listening, our observation about the distinction between how people listen to nostalgic and catalog music is relevant to studies of music, culture, and memory more generally. It may, for example, motivate new approaches to catalog sampling for surveys and experiments. Qualitative work that can better account for listener context and perceptions will continue to be vital for explaining why people listen to the music they do.

Our analysis compared patterns of nostalgic listening across four national contexts. The findings for each country are numerically different, but the general trends tend to be consistent for each result. Adding to Hepper et al. (2014)'s finding that nostalgia is a pancultural emotion, this suggests cross-cultural similarity in how people use nostalgic music in their daily life. However, we found that the music people label nostalgic may vary by a person's socio-linguistic context. Respondents from Spanish-speaking countries tended to label sadder-sounding music nostalgic, and respondents in English-speaking countries tended to label happier-sounding music nostalgic. Research that better accounts for the connotation of nostalgia and related emotions (e.g., saudade) in different socio-linguistic contexts may find greater differences in listening behavior across national contexts.

There are several important limitations to our study. First, it is based on music streaming data. This could affect our findings in a few ways. We do not capture music listening via other media, such as physical media (CDs, records), radio, or other streaming services. However, we restricted our sample to consistent users to limit inclusion of people whose primary means of listening to music is not Spotify. Use of streaming services is correlated with age (Krumhansl 2017). There is likely selection bias among the listeners we studied, especially among older listeners.

A second limitation is that our analysis of nostalgic listening is based on estimates made by a classifier. To oversample nostalgic music, we used a classifier developed with a survey of employees. Doing so may introduce bias to the classifier, especially if the features we used or the opinions of employees are very different from concepts of nostalgia among Spotify's global listeners. The classifier also treats nostalgic music as a stable category, though the emotions people associate with music may vary over time (Holland and Kensinger 2010; Kensinger and Ford 2020). Our final classifier had a precision of 0.71 and a recall of 0.48. This means our analysis of nostalgic listening includes some songs that are not nostalgic to listeners, and omits many songs that are. Although imperfect, we were surprised by the classifier's accuracy given that what makes a song nostalgic may be affected by memory, past experience, and past taste we could not directly observe.

A major strength of this research is that it offers a unique picture of nostalgia by drawing on a large observational dataset. How well studies that use experiments and surveys generalize to everyday life is disputed (Newman et al. 2020; Jakubowski and Ghosh 2021). We analyze nostalgic listening across countries, age groups, and time. This helps to validate past findings, and helps to illustrate methodological considerations for studies based on aggregated data or that require sampling the historical music catalog.

In addition to the suggestions outlined above, future research can build on this work by continuing to explore the relationship between personal memories associated with music and collective memory of music (e.g. van Dijck 2006; Spivack et al. 2019; Krumhansl and Zupnick 2013). Our work shows many differences in how people interact with nostalgic music compared to how they interact with catalog music generally. Yet "nostalgia" is sometimes used to refer to the act of revisiting any historical artifact. Understanding how collective ideas about the past are affected by personal feelings of nostalgia, and how personal feelings of nostalgia are affected by collective memories of the past can clarify the connection between the stories we tell about ourselves and the stories we tell about our societies.

Appendix

Appendix Figure A1: Pool of Potentially nostalgic music by birth year Appendix Figure A1 is based on streaming in 2016. For listeners in each birth year, we summed the total number of streams to tracks released during that group's child and young adulthood and divided it by their

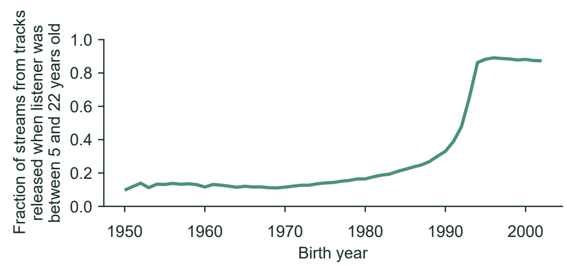

Figure 9: Fraction of streams to tracks released when listeners were between 5–22 years old for each age group. Based on streaming data from 2016.

total streams. We define child and young adulthood as between about age 5–22, the period of music typically associated with nostalgic feelings. Among listeners born in 1985 or earlier (age 31+ at the time of this data), streams to music from the child and young adulthood accounts for 20% or less of their listening. The proportion increases for respondents closer to their child and young adulthood. Note that in the graph above, listeners born in 1994 or earlier are still in their young adulthood. When we sampled listeners' streaming histories to create assumed negative data for our classifier, we did so for the years 2016 to mid-2021. Therefore, the figure above shows the upper-bound for the fraction of streams to one's childhood and young adulthood for the young listeners in our sample. Nostalgic music is a subset of music released during one's child and young adulthood, which we expect to be relatively small.

Country	Respondents	Responses	% Positive
United States	3583	39306	65.2
Brazil	3271	35309	67.2
United Kingdom	2623	30218	64.1
Mexico	2314	26129	53.1
Germany	1427	16650	59.8
Spain	1167	13561	56.6
Canada	1008	10817	64.9
Argentina	874	9835	52.2
Australia	639	7187	65.7
Philippines	454	4916	70.3
Singapore	187	2083	70.4

Table A1: Survey statistics for each country, including number of unique respondents and responses, as well as the percentage of responses indicating a nostalgic song.

Appendix Figure A2: Peaks in Catalog Listening by Nostalgic Labeling, Cohort Over-Indexing, and Classifier Output Appendix Figure A2 shows that the period of music respondents most often labeled as nostalgic is different from the period of music respondents disproportionately listen to. This suggests the music respondents stream disproportionately is different from the music that most triggers nostalgic feeling. This is especially interesting because research about autobiographical memory has consis-

Peaks in Catalog Listening by Nostalgic Labeling, Cohort Over-Indexing, and Classifier Output

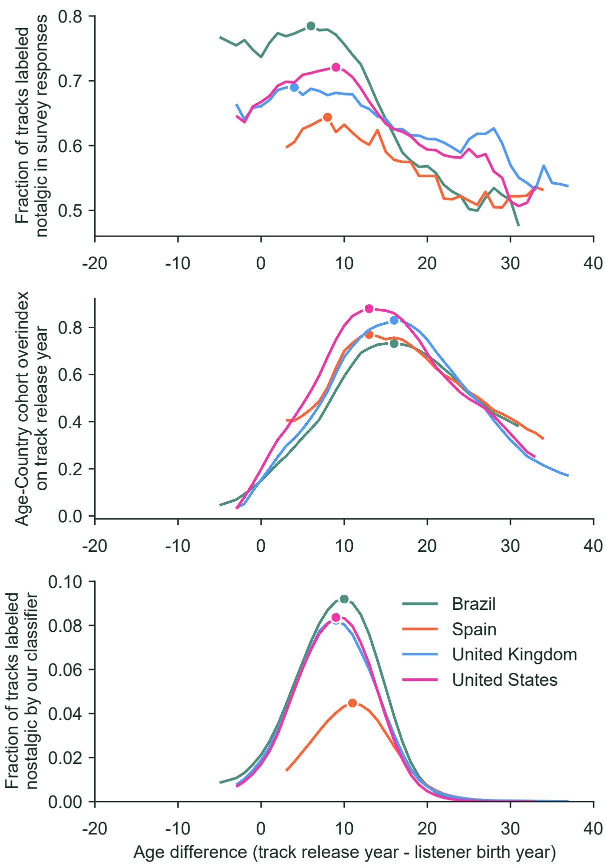

Figure A1: The top panel shows the fraction of tracks survey respondents labeled nostalgic. The middle panel shows how much respondents from the same country and age group disproportionately listen to music released at a given age. The bottom panel shows the fraction of tracks our classifier labeled nostalgic. Points indicate the peak of the distribution.

tently found evidence of a "reminiscence bump" — the phenomenon that people tend to have more and more vivid memories from their teen and young adult years than from other periods in their lives (Munawar, Kuhn, and Haque 2018). That labels of nostalgic music peak on tracks released when respondents were very young could underscore how nostalgic feeling is not just about memory, but also about the feelings and idealization associated with memories. However, we did not sample tracks evenly or randomly across respondents' lives. If our method tended to select more nostalgic music for the periods we sampled fewer tracks from, selection bias rather than nostalgic emotion would explain the early peak shown in the top panel. Further research is needed to understand whether and why the reminiscence bump and artifacts that trigger nostalgic emotion tend to come from different periods.

Appendix Figure A3: Difference in Valence Among

Difference in Valence Among Tracks Labeled Nostalgic versus Not Nostalgic by Country

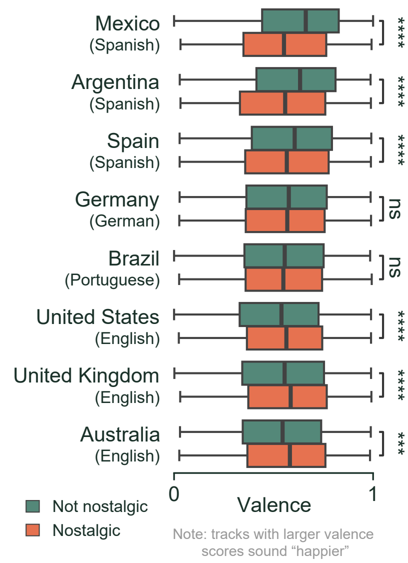

Figure A2: The figure above shows the distribution of valence among tracks labeled nostalgic or not among survey respondents. *significant at $p<0.05$, **significant at $p<0.01$, ***significant at $p<0.001$, ****significant at $p<0.0001$.

Tracks Labeled Nostalgic versus Not Nostalgic by Country Appendix Figure A3 shows the distribution of valence among tracks labeled nostalgic and not nostalgic among survey respondents. Valence is measured with a widely used model-based measure designed to distinguish happy from sad-sounding music using an acoustic waveform (The Echo Nest Blog 2013). Valence scores were significantly lower for tracks labeled nostalgic among listeners in Spanish-speaking countries[1]. Valence scores were significantly higher for tracks labeled nostalgic among listeners in English-speaking countries[2]. Valence scores were not significantly different for respondents in Brazil t(35,308)=0.9562, p=1.000) or Germany t(16,649)=1.156, p=1.000). All tests are two-tailed and reported with the Bonferroni correction to avoid erroneous rejection of the null hypothesis.

This trend may explain why our classifier had lower performance for respondents in Spain than in the other countries in our analysis. The difference in valence among countries is consistent with the connotations of the word nostalgia in different cultures. While "nostalgia" generally holds a positive connotation in English, in Spanish "nostálgica" is more associated with feelings of sadness and homesickness (Cambridge Dictionary 2021). We have less data from respondents in Spanish-speaking countries than English speaking

[1]Spain: t(13,560)=5.965, p=0.000; Mexico: t(26,128)=22.39, p=0.000; Argentina: t(9,834)=11.21, p=0.000)

[2]United States: t(39,305)=-10.89, p=0.000; United Kingdom: t(30,217)=-7.735, p=0.000; Australia: t(7,186)=-4.203, p=0.0002)

countries, and the classifier trained to sample data was developed with respondents in the US and Sweden. Though listener country and valence are features in the classifier, we may not have the data necessary for the classifier to replicate "nostálgica" as well as it replicates "nostalgia."

Broader perspective, ethics and competing interests We discuss two areas of broader impact of our work: impacts on human subjects, and impacts on future research.

We believe our research poses minimal risk to the people included in our analysis. Respondents to our listener survey participated voluntarily and understood they were being surveyed by Spotify for research purposes. Our observational analysis presents no further risk to participants than what they would encounter from use of digital services in their daily life. All findings are aggregate statistics based on large samples. We will not release the raw data on which this analysis is based. While this makes our research less transparent and precludes replication of our study, releasing the raw data poses risk to subjects of de-anonymization, release of sensitive information, and violation of privacy expectations.

Our work may also impact future research on nostalgia, music listening, emotion, and memory. Researchers are rarely able to observe the behavior of large groups at the level of individual people over such a long period of time. This analysis provides context for estimates of behavior based on other methods. However, we make several assumptions to conduct this study. Most notably, that our sample is generalizable and that our classifier is relatively accurate. We discuss why we believe these assumptions are reasonable throughout the paper. However, if our assumptions are incorrect, our findings would be misleading. We discuss the limitations of our work throughout the paper and note areas that would especially benefit from additional research.

Entropy of nostalgic streams by birth year

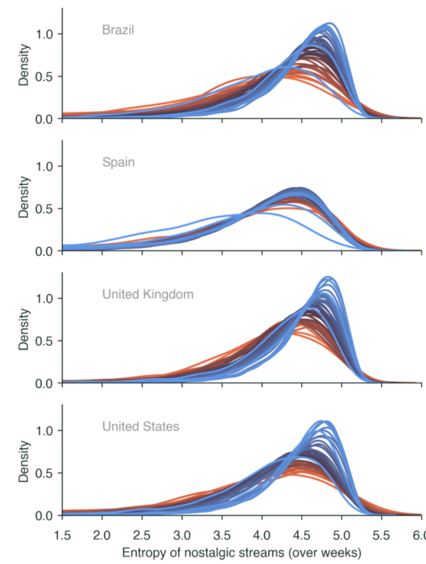

Figure A3: The figure above shows the distribution of entropy of streams of nostalgic content over weeks of the 5.5 year sample by birth year. The youngest listeners are in the lightest blue, while the oldest listeners are in the brightest orange. Higher scores indicate more consistent nostalgic streaming, and lower scores indicate nostalgic streams tend to be more clustered in time. The figure indicates most people have relatively high entropy scores overall. However, younger listeners tend to have higher entropy scores than older listeners in all of the contexts but Spain.

References

Abakoumkin, G.; Wildschut, T.; and Sedikides, C. 2020. Nostalgia proneness and the collective self. *Frontiers in Psychology*, 11: 2829.

Batcho, K. I. 2013. Nostalgia: The bittersweet history of a psychological concept. *History of Psychology*, 16(3): 165.

Bonneville-Roussy, A.; Rentfrow, P. J.; Xu, M. K.; and Potter, J. 2013. Music through the ages: Trends in musical engagement and preferences from adolescence through middle adulthood. *Journal of Personality and Social Psychology*, 105(4): 703.

Cabedo-Mas, A.; Arriaga-Sanz, C.; and Moliner-Miravet, L. 2020. Uses and perceptions of music in times of COVID-19: a Spanish population survey. *Frontiers in Psychology*, 11.

Cambridge Dictionary. 2021. nostalgia. https://dictionary.cambridge.org/dictionary/spanish-english/nostalgia. Accessed: 2022-04-18.

Davis, F. 1979. *Yearning for yesterday: A sociology of nostalgia.* Free Press.

De Diego, F. F.; Ots, C. V.; et al. 2014. Nostalgia: a conceptual history. *History of Psychiatry*, 25(4): 404–411.

DeNora, T. 2000. *Music in everyday life*. Cambridge University Press.

Farese, G.; and Asano-Cavanagh, Y. 2019. Analysing nostalgia in cross-linguistic perspective. *Philology*, 4(2018–2019): 213–241.

Fink, L. K.; Warrenburg, L. A.; Howlin, C.; Randall, W. M.; Hansen, N. C.; and Wald-Fuhrmann, M. 2021. Viral tunes: changes in musical behaviours and interest in coronamusic predict socio-emotional coping during COVID-19 lockdown. *Humanities and Social Sciences Communications*, 8(180).

Gibbs, H.; and Egermann, H. 2021. Music-Evoked Nostalgia and Wellbeing During the United Kingdom COVID-19 Pandemic: Content, Subjective Effects, and Function. *Frontiers in Psychology*, 12: 787.

González, A. 2012. Adiós a la nostalgia: la narrativa hispanoamericana después de la nación. *Revista de Estudios Hispánicos*, 46(1): 83–97.

Granot, R.; Spitz, D. H.; Cherki, B. R.; Loui, P.; Timmers, R.; Schaefer, R. S.; Vuoskoski, J. K.; Cárdenas-Soler, R.-N.; Soares-Quadros Jr, J. F.; Li, S.; et al. 2021. "Help! I Need Somebody": Music as a Global Resource for Obtaining Wellbeing Goals in Times of Crisis. *Frontiers in psychology*, 12: 1038.

Hepper, E.; Wildschut, T.; Sedikides, C.; Ritchie, T.; Yung, Y.; Hansen, N.; Abakoumkin, G.; Arikan, G.; Cisek, S.; Demassosso, D.; Gebauer, J.; Gerber, J.; González, R.; Kusumi, T.; Misra, G.; Rusu, M.; Ryan, O.; Stephan, E.; Vingerhoets, A.; and Zhou, X. 2014. Pancultural nostalgia: prototypical conceptions across cultures. *Emotion*, 14(4): 733–47.

Hepper, E. G.; Wildschut, T.; Sedikides, C.; Robertson, S.; and Routledge, C. D. 2021. Time capsule: Nostalgia shields psychological wellbeing from limited time horizons. *Emotion*, 21(3): 644–664.

Hird, E.; and North, A. 2021. The relationship between uses of music, musical taste, age, and life goals. *Psychology of Music*, 49(4): 872–889.

Hofstede, G. 1984. *Culture's consequences: International differences in work-related values*, volume 5. Sage.

Holbrook, M. 1993. Nostalgia and Consumption Preferences: Some Emerging Patterns of Consumer Tastes. *Journal of Consumer Research*, 20(2): 245.

Holland, A. C.; and Kensinger, E. A. 2010. Emotion and autobiographical memory. *Physics of Life Reviews*, 7(1): 88–131.

Jakubowski, K.; and Ghosh, A. 2021. Music-evoked autobiographical memories in everyday life. *Psychology of Music*, 49(3): 649–666.

Juhl, J.; Wildschut, T.; Sedikides, C.; Xiong, X.; and Zhou, X. 2021. Nostalgia promotes help seeking by fostering social connectedness. *Emotion*, 21(3): 631–643.

Kensinger, E. A.; and Ford, J. H. 2020. Retrieval of emotional events from memory. *Annual Review of Psychology*, 71: 251–272.

Krumhansl, C. L. 2017. Listening Niches across a Century of Popular Music. *Frontiers in Psychology*, 8: 431.

Krumhansl, C. L.; and Zupnick, J. A. 2013. Cascading Reminiscence Bumps in Popular Music. *Psychological Science*, 24(10): 2057–2068.

Merchant, A.; and Rose, G. M. 2013. Effects of advertising-evoked vicarious nostalgia on brand heritage. *Journal of Business Research*, 66: 12.

Munawar, K.; Kuhn, S. K.; and Haque, S. 2018. Understanding the reminiscence bump: A systematic review. *PLOS One*, 13(12): 1–36.

Newman, D. B.; Sachs, M. E.; Stone, A. A.; and Schwarz, N. 2020. Nostalgia and Well-Being in Daily Life: An Ecological Validity Perspective. *Journal of Personality and Social Psychology*, 118: 2.

Park, M.; Thom, J.; Mennicken, S.; Cramer, H.; and Macy, M. 2019. Global music streaming data reveal diurnal and seasonal patterns of affective preference. *Nature Human Behaviour*, 3: 230–236.

Routledge, C.; Arndt, J.; Wildschut, T.; Sedikides, C.; Hart, C. M.; Juhl, J.; Vinger-hoets, A. J. J. M.; and Schlotz, W. 2011. The past makes the present meaningful: Nostalgia as an existential resource. *Journal of Personality and Social Psychology*, 101(3): 638–652.

Roy, W. G.; and Dowd, T. J. 2010. What Is Sociological about Music? *Annual Review of Sociology*, 36(1): 183–203.

Sanna Passino, F.; Maystre, L.; Moor, D.; Anderson, A.; and Lalmas, M. 2021. *Where To Next? A Dynamic Model of User Preferences*, 3210–3220. New York, NY, USA: Association for Computing Machinery.

Schlagman, S.; Kliegel, M.; Schulz, J.; and Kvavilashvili, L. 2009. Differential effects of age on involuntary and voluntary autobiographical memory. *Psychology and Aging*, 24(2): 397.

Sedikides, C.; and Wildschut, T. 2018. Finding Meaning in Nostalgia. *Review of General Psychology*, 22(1): 48–61.

Sedikides, C.; Wildschut, T.; and Baden, D. 2004. Nostalgia: Conceptual Issues and Existential Functions. In *Handbook of Experimental Existential Psychology*, 200–214. Guilford Press.

Spivack, S.; Philibotte, S. J.; Spilka, N. H.; Passman, I. J.; and Wallisch, P. 2019. Who remembers the Beatles? The collective memory for popular music. *PLOS One*, 14(2): e0210066.

Taylor, S. J.; and Letham, B. 2018. Forecasting at Scale. *The American Statistician*, 72(1): 37–45.

The Echo Nest Blog. 2013. Plotting Music's Emotional Valence, 1950-2013.

Triandis, H. C. 1995. *Individualism and collectivism*. Routledge.

van Dijck, J. 2006. Record and Hold: Popular Music between Personal and Collective Memory. *Critical Studies in Media Communication*, 23(5): 357–374.

van Tilburg, W. A. P.; Igou, E. R.; and Sedikides, C. 2013. In search of meaningfulness: Nostalgia as an antidote to boredom. *Emotion*, 13(3): 450–461.

Vidas, D.; Larwood, J. L.; Nelson, N. L.; and Dingle, G. A. 2021. Music Listening as a Strategy for Managing COVID-19 Stress in First-Year University Students. *Frontiers in Psychology*, 12: 922.

Weinberger, M. F. 2015. Dominant Consumption Rituals and Intragroup Boundary Work: How Non-Celebrants Manage Conflicting Relational and Identity Goals. *Journal of Consumer Research*, 42(3): 378–400.

Yang, Z.; Sedikides, C.; Izuma, K.; Wildschut, T.; Kashima, E. S.; Luo, Y. L.; Chen, J.; and Cai, H. 2021. Nostalgia enhances detection of death threat: neural and behavioral evidence. *Scientific Reports*, 11(1): 1–8.

Yeung, T. Y.-C. 2020. Did the COVID-19 Pandemic Trigger Nostalgia? Evidence of Music Consumption on Spotify. *SSRN Electronic Journal*.

Zhou, X.; Wildschut, T.; Sedikides, C.; Shi, K.; and Feng, C. 2011. Nostalgia: The Gift That Keeps on Giving. *Journal of Consumer Research*, 39(1): 39–50.